ARI Research Note 92-36

Empirical Development of a Scale for the Prediction of Performance on a Sustained Monitoring Task

Linda G. Pierce and Lloyd M. Crumley
U.S. Army Research Institute

Field Unit at Fort Sill, Oklahoma
Edwin R. Smootz, Acting Chief

MANPRINT Division
Robin L. Keesee, Director

May 1992

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

MICHAEL D. SHALER
COL, AR
Commanding

Technical review by

James S. Ainsworth

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

1992, May	Final	Jun 91 - Jan 92

Empirical Development of a Scale for the Prediction of Performance on a Sustained Monitoring Task

62785A
791
1212
H01

Pierce, Linda G.; and Crumley, Lloyd M.

U.S. Army Research Institute for the Behavioral and Social Sciences
Fort Sill Field Unit
Fort Sill, OK 73503-5600

ARI Research Note 92-36

--

--

--

Approved for public release; distribution is unlimited.

--

Research on vigilance performance has been extensive, but few guidelines exist for selecting persons well suited to perform vigilance tasks. The focus of this pilot study was to determine if a subset of items from the Minnesota Multiphasic Personality Inventory (MMPI) that would discriminate between soldiers able to sustain an adequate level of performance and soldiers whose performance deteriorated over time could be identified. Thirty-nine items produced chi-squares with p values less than .05. A discriminant analysis correctly classified 88.5% of the soldiers maintaining their performance rate and 83.9% of the soldiers with performance degraded over time. If validation efforts are successful, it may be possible to use the 39-item scale to select soldiers who will, on average, detect 13% more targets and incorrectly identify 15% fewer targets.

Target detection
Perception

Scale development
Selection

28

--

Unclassified

Unclassified

Unclassified

Unlimited

EMPIRICAL DEVELOPMENT OF A SCALE FOR THE PREDICTION OF PERFORMANCE ON A SUSTAINED MONITORING TASK

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Empirical Development of a Scale for the Prediction of Performance on a Sustained Monitoring Task

INTRODUCTION

In 1962, Dobbins and Skordahl reported, based on Army-wide task analyses, that at least some monitoring activity was performed by soldiers carrying out over 100 of the 1500 Army jobs analyzed (cited in Wiener, 1987). That number has increased and will continue to increase as the degree of automation within the Army continues to grow.

A weapon system in which monitoring behavior is necessary for successful mission performance is a remotely piloted vehicle (RPV). RPV's fly over enemy terrain photographing and transmitting a continuous, real-time video from an on-board camera. The camera is controlled from a ground station where a crew member, the Sensor Station Operator (SSO), observes the terrain picture on a screen as it is transmitted from the RPV. A major task for the SSO is to search the TV screen and take action when a target appears. Missions can involve flights of two or more hours. During this period, SSO's attend carefully to the screen, remain alert, monitor and search the screen for potential targets, and recognize as targets certain objects that appear infrequently and are often only poorly visible. This type of attention-demanding task can be classified as a complex monitoring task with high spatial and temporal uncertainty (Parasuraman, 1986).

Most of the research that relates to how well people perform tasks that require extended periods of attention in situations that are not conducive to such behavior has been conducted in a related area--vigilance. Monitoring tasks are related to, yet distinguished from, vigilance tasks. Monitoring tasks are generally considered those tasks in which the signals are continuous and rather complex. In performing a monitoring task the "observer must attend actively to a source or many sources of stimuli to identify a previously specified signal, or infer the presence of a 'signal,' or estimate the 'parameters' of some process from data presented on the sources" (Parasuraman, 1986, p. 43-2). The term vigilance was originally coined by Head (1923) and was later defined by N.H. Mackworth (1957), as "... a state of readiness to detect and respond to certain specified small changes occurring at random time intervals in the environment" (p. 389).

During World War II, N.H. Mackworth (1950) studied vigilance behavior in order to determine the optimum length of watch for sonar and radar operators on antisubmarine patrol. Since that time, vigilance researchers have been among the most copious behavioral scientists, publishing over 1000 studies (Mackie, 1987). Although the literature is extensive, researchers have

tended to focus on explaining the inability of individuals to maintain vigilant behavior over extended periods of time. This phenomenon, identified and studied by N.H. Mackworth (1950), is commonly referred to as the vigilance decrement (Davies & Parasuraman, 1982). Many researchers have reported large individual differences in susceptibility to the vigilance decrement and have proposed explanations which emphasize both task-related variables, such as signal characteristics (e.g., modality, strength, frequency, and probability) and individual observer characteristics; such as sex, age, intelligence, motivation, and personality or temperament (see Boff & Lincoln, 1988 for a review). Despite the amount of research which has been conducted in the area, Mackie, in a 1987 review article, concluded that relatively little is known regarding ways to improve vigilance performance.

Much of the little that is known about improving vigilance, or monitoring behavior, results from carefully controlled laboratory studies, which have not been validated in an operational environment. Furthermore, a majority of these studies were designed to assess the effect of task-related variables on the vigilance decrement. Indeed, of the 18 strategies identified by either Davies and Parasuraman (1982) or Craig (1984; cited in Mackie, 1987) for improving vigilance behavior, only one dealt with selection and it was too general to be useful. The general strategy was to "use personnel selection techniques to identify individuals with a propensity for maintaining vigilance" (cited in Mackie, 1987, p. 710). In contrast, suggestions for altering the task to improve vigilance performance were quite specific (e.g., reduce temporal and spatial uncertainty of signal appearance, increase signal "conspicuity" [sic], and provide performance feedback).

In many situations, for example the RPV-SSO position, it is not possible to alter the monitoring task, hence, the most viable strategy remaining is to use personnel selection techniques to identify individuals who will perform better in the task either because they have more aptitude for the task or are better suited by temperament to attend to the task. Selection of higher aptitude operators who are also likely to remain attentive to the task over time may be possible if it can be shown that individuals who vary in their ability to perform such tasks can be identified on the basis of simpler tasks, such as an aptitude test or a temperament scale. Thus, there are two problems: (a) identification of those individuals who are best able to see targets, and (b) identification of those individuals who are better able to maintain their level of performance throughout the mission. The first issue, aptitude, was addressed in earlier research at the U.S. Army Research Institute (Crumley, Pierce, Schwalm, Coke, & Brown, in preparation). The researchers found that the cognitive factor Speed of Closure, as measured by the Snowy Pictures Test, the Concealed Words Test, and the

Gestalt Completion Test does predict better target detection performance. The correlations, though weak, were significant (r = .263, p < .01; r = .134, p < .05; r = .145, p < .01; and a mean correlation for the three tests of r = .181, p < .05).

The research, discussed in the present paper, addresses the second issue--temperament. We sought to determine if individuals who are best able to see targets and maintain their performance throughout the mission could be identified on the basis of temperament. In the research reported here, we administered the Minnesota Multiphasic Personality Inventory (MMPI) to a large group of subjects who were then tested in a situation which simulated the SSO monitoring task. Our intent was to see if a pool of MMPI items could be identified and, with proper validation, used to create a short, easy-to-administer scale that would predict the persons most capable of sustaining attentiveness over a long monitoring period. Since the development of the MMPI, researchers have used the MMPI items to construct numerous special personality scales (see Barron, 1953; Taylor, 1951). More recently, researchers have attempted to use the MMPI to develop special personality scales in diverse areas as predicting athletic potential (Brown, Morgan, & Kihlstrom, 1989) and post-operative outcome (Sorensen & Mors, 1988), and detecting early sexual abuse (Roland, Zelhart, & Dubes, 1989).

The MMPI has also been used by the military for a variety of purposes. Butcher et al. (1990) reported that the military services have used the MMPI (a) for selection of military personnel for special duty assignments, (b) to evaluate and to predict training failures, (c) to assess the impact of harsh environmental conditions or imprisonment in prisoner-of-war camps, and, perhaps most commonly, (d) in military medical centers and correctional facilities to assess a wide assortment of clinical problems. For example, Fairbank, Keane, and Malloy (1983) used the MMPI to detect the symptoms of combat-related posttraumatic stress disorder (PTSD) in Vietnam-era veterans. They found that the veterans with a PTSD diagnosis scored significantly higher than the well-adjusted group on all the clinical scales except the Masculinity-Femininity scale, and that there were significant differences between the PTSD group and the veterans with psychological problems other than PTSD on the Hypochondriasis, Hysteria, and Psychasthenia scales.

Later Keane, Malloy, and Fairbank (1984) expanded this work and looked at the responses of male veterans to specific items on the MMPI. Their subjects were 100 combat veterans with a diagnosis of PTSD and 100 veterans with a diagnosis other than PTSD. Once again, the PTSD group scored significantly higher on all clinical scales, except the Masculinity-Femininity scale, replicating the earlier results of Fairbank et al. (1983). The MMPI items were then submitted to chi-square analyses. Forty-nine items were found to be endorsed differentially by the two

groups with p values less than .001. The PTSD scale that resulted correctly classified 82% of both samples. An attempt to cross-validate the PTSD subscale in a separate population of veterans has been somewhat successful. Gayton, Burchstead, and Matthews (1986) were able to correctly classify 46% of the PTSD group and 78% of the non-PTSD group after controlling for potential misdiagnoses.

As stated previously, the purpose of the present study is to identify a subset of MMPI items to create a scale which could be used to identify persons who are likely to maintain their best level of monitoring performance over extended periods of time. The target detection task will be performed in a laboratory in which a simulation of the RPV operator task has been created. Effective target detection will be defined as the ability of soldiers to initially identify a relatively high number of targets and then to be able to maintain or improve on that initial performance over time. Thus, we will attempt to identify a subset of MMPI items that will allow the Army to select the best soldiers to perform monitoring tasks who are the least likely to evidence the classic vigilance decrement.

METHOD

Subjects

A total of 240 enlisted soldiers awaiting reassignment after completing training at the U.S. Army Field Artillery School were tested. Data on some of the subjects were incomplete. Gaps in the data were due to occasional equipment malfunctions or to subjects' not being available for portions of the testing because of transfer to duty assignments. Data from 174 subjects were used in the present study. There were 161 males and 13 females with a mean age of 19.64. Data collected on each subject included scores on the MMPI Form R, ASVAB subtest scores, scores on selected cognitive factor tests, a number of biographic items, and two related measures of performance--targets detected and false detections or false alarms. The data used in the present analyses include the MMPI item scores and the measures of performance. Analyses, using the other data collected are reported in Crumley et al. (in preparation).

Apparatus

Subjects were tested in groups of four in an air-conditioned 15' x 18' room. Windows in the room were masked, and the room was darkened during the data collection process. The only light sources were a lamp on the desk of the experimenter and a single-tubed, ceiling-mounted fluorescent light. Both light sources were behind the subjects. Subjects were seated behind tables that were 10'6" from the center of a 19" TV screen. Each subject was provided a push-button switch, mounted on a small block of

wood which could be moved to whatever position was considered comfortable. The subject depressed the switch when he or she saw, or thought they saw, a target on the screen. Switch closures were recorded by an event recorder (Lafayette Instrument Company, Model 56042, Six Channel Mini Recorder) which also recorded when a target was present from a signal placed on the video tape audio track.

The stimuli seen by the subjects consisted of a video presentation of 1440 slides showing aerial views of terrain on and near Fort Sill, OK. The original photographs were taken with a 35mm camera from a low-flying helicopter whose altitude varied between 400 and 800 feet. Photographs were taken as the helicopter approached, passed over, and flew beyond the prepositioned targets. As the photographs were taken, the pilot tried to adjust his speed so that each photograph showed a terrain section that only slightly overlapped the previous picture. When the pilot was successful the target appeared on only one slide in the sequence. If the speed was low the overlap was greater and a target could appear twice or three times in a series. When this occurred, two or three target series were created by taking alternate slides, or every third slide, from the set to form multiple series with the same target.

Targets photographed were trucks, jeeps, buildings, bridges, and helicopters on the ground. Targets were hidden to varying degrees by their placement in the fields and woods, and on the roads. From the pool of slides that contained targets, 90 target series were selected for inclusion in the test sequence. The researchers based the selections on the following criteria: (a) a good transition between target series had to be possible, (b) only one target should appear on a slide, (c) a target should not be so obscure that no subject would be likely to see it or so obvious that everyone would probably see it.

After the slides were examined, they were mounted in a carrousel projector and displayed on a rear-projection screen. A video camera mounted facing the front of the screen was then used to create the three 40-minute tapes used in the experiment. Because the 4/3 aspect ratio of the video screen was smaller than the aspect ratio of the 35mm pictures, the location of targets and content of the individual views could be controlled somewhat by adjusting the projected image left or right. This feature was used to position the target further left or right, and to avoid situations where some feature near the edge of the film was detrimental to the appearance of the terrain view. The final criterion--target detectability--was assured by having the researchers examine the black and white video display of the slides containing the targets to be certain that each target was indeed visible.

Procedure

Testing began on Monday of each week when the subjects, in groups of 20, were asked to complete a biographic inventory, then given the MMPI Form R and selected cognitive style tests. Later the same week, Tuesday morning through Friday afternoon, subjects, in groups of four, returned and performed the monitoring task. Prior to the target detection test, the subjects were read instructions which described the nature of the test, the types of targets, and the purpose of the research.

Before the data collection process began, everyone was shown two training tapes. First, the subjects saw a 16-frame video tape containing targets, and the experimenter pointed out the location and type of target as the frames passed. Subjects were instructed to report as targets any man-made object other than roads, railroads, power lines, and fences; and were told that targets were considered to be such things as vehicles, buildings of any type, and bridges. Subjects then were shown a six-and-one-half-minute tape containing 80 slides of which 27 had targets. As the tape was played, subjects depressed their switches when they saw a target. The same tape was then played with a pointer overlay showing the location of each of the targets. The experimenter then answered questions, after which the subjects took a brief break.

The test process took two hours. Each of the slides remained visible for four-and-one-half seconds and there was a one-half-second dark period between frames. The target detection task was divided into three blocks, each lasting 40 minutes. Blocks were separated by a 10-second period, which was used to switch video tapes. There were 480 stimuli per block with 25, 37, and 28 targets, respectively. Thus, 6.25% (90) of the 1440 frames contained targets and 1350 did not contain a target.

Targets identified correctly were called correct detections or hits. Beginning with the experiments conducted by Mackworth (1948, 1950), this has been the most frequently reported response measure, and, typically, as was done here, the task is divided into successive time periods and the hit rate computed for each period. However, vigilance performance over time cannot be based solely on the detection rate of targets, but must also include the rate of incorrect detections or false alarm rate. Errors of commission or false alarms occur when the observer reports a signal when none has occurred. Thus, both the hit rate and the false alarm rate were calculated. The hit rate was computed by dividing the number of correct detections by the sum of correct detections and misses to arrive at an average hit rate for each interval. The false alarm rate was calculated by dividing the number of responses to nontargets by the number of nontargets for each recording period. For computations, both hits and false alarms were then transformed into percentages.

RESULTS

Target Detection

Detection scores were computed for each subject by subtracting the percentage of incorrect detections (false alarms) from the percentage of correct detections (hits) for each block. Presented in Table 1, for each of the three blocks, are the means and standard deviations for the percentage of: (a) correct detections, (b) false alarms, and (c) correct detections minus false alarms.

Table 1

Means and Standard Deviations for the Percentage of Correct Detections, False Alarms, and Correct Detections Minus False Alarms

		Blocks		
		1	2	3
Correct Detections				
	Mean	33.84	27.35	41.25
	SD	15.35	15.78	21.13
False Alarms				
	Mean	9.80	7.98	8.88
	SD	6.42	5.95	7.81
Correct Detections Minus False Alarms				
	Mean	24.03	19.37	32.37
	SD	13.00	14.10	19.91

Note. n=174

Stanine scores were then computed for each subject by block. Based on their stanine scores, subjects were divided into two groups. One group consisted of those persons who met the following three performance criteria: (a) their Block 1 performance position was stanine 3 or higher; (b) their Block 2

position was either higher, equal, or not more than one stanine below their Block 1 performance position; and (c) their Block 3 performance was either higher, equal, or not more than one stanine below their Block 2 performance and not more than one stanine below their Block 1 performance. There were 87 subjects whose scores on Block 1 were stanine three or higher, and whose subsequent Block scores did not decrease more than one stanine. These subjects were assigned to Group 1 (Maintainers), and the remaining 87 subjects to Group 2 (Nonmaintainers).

Operationally, both vigilance decrement and level of vigilance are important considerations (Parasuraman, 1986). Thus, the performance maintainers were chosen in this manner to exclude (a) soldiers whose performance was satisfactory to begin with but then declined and (b) soldiers who did not evidence a vigilance decrement but whose initial performance in Block 1 was so low that it could not decrease enough to make them excludable on the basis of a two-stanine decrement.

Data for the Maintainers, Nonmaintainers, and for the groups combined are shown in Table 2. As can be seen, all subjects tended to show a decrement during the middle portion of the test period. Despite this, our criteria for assigning subjects to the Maintainer group ensured that we would select subjects who performed the most consistently over time. Because we computed stanines, the degree of consistency shown by each group was relative to the performance of all the subjects in each block.

An analysis of variance was computed, and performance, as measured by the number of correct detections minus the number of false alarms, was found to differ significantly between vigilance groups (Maintainers and Nonmaintainers) and across blocks (see Table 3). Maintainers ($\underline{M}$ = 30.14, $\underline{SD}$ = 13.35), as expected, performed significantly better than Nonmaintainers ($\underline{M}$ = 19.03, $\underline{SD}$ = 11.34) and, as would be predicted from previous vigilance studies, performance decreased significantly between Blocks 1 and 2, and then improved significantly between Blocks 2 and 3 (see Davies & Parasuraman, 1982 for a review on the vigilance decrement). The difference in performance between Blocks 1 and 3 was also significant. There was also a significant interaction between vigilance groups across blocks (see Figure 1).

The two performance groups scored similarly on Block 1, somewhat differently on Block 2, and very differently on Block 3. Both Block 2 and Block 3 differences were statistically significant (see Figure 1). This was, as expected, based on the criteria for inclusion into the Maintainer Group.

Table 2

Means and Standard Deviations for the Percentage of Correct Detections Minus False Alarms by Performance Group

		Blocks		
Group		1	2	3
Maintainers				
	Mean	25.24	24.34	42.46
	SD	11.87	14.19	17.74
Nonmaintainers				
	Mean	22.83	14.40	22.29
	SD	14.00	12.20	16.63
Total Sample				
	Mean	24.03	19.37	32.38
	SD	13.00	14.10	19.91

Table 3

Analysis of Variance Summary Table for Performance Group Across Performance Blocks

Source	Sum of squares	Degrees of freedom	Mean square	F
Between subjects				
Vigilance Group	15336.90	1	15336.90	33.17*
Error	79524.20	172	462.35	
Within subjects				
Trials	15108.04	2	7554.02	85.36*
Trials x Group	6915.08	2	3457.54	39.07*
Error	30444.00	344	88.50	
TOTAL	147328.22	521		

* $p < .00001$

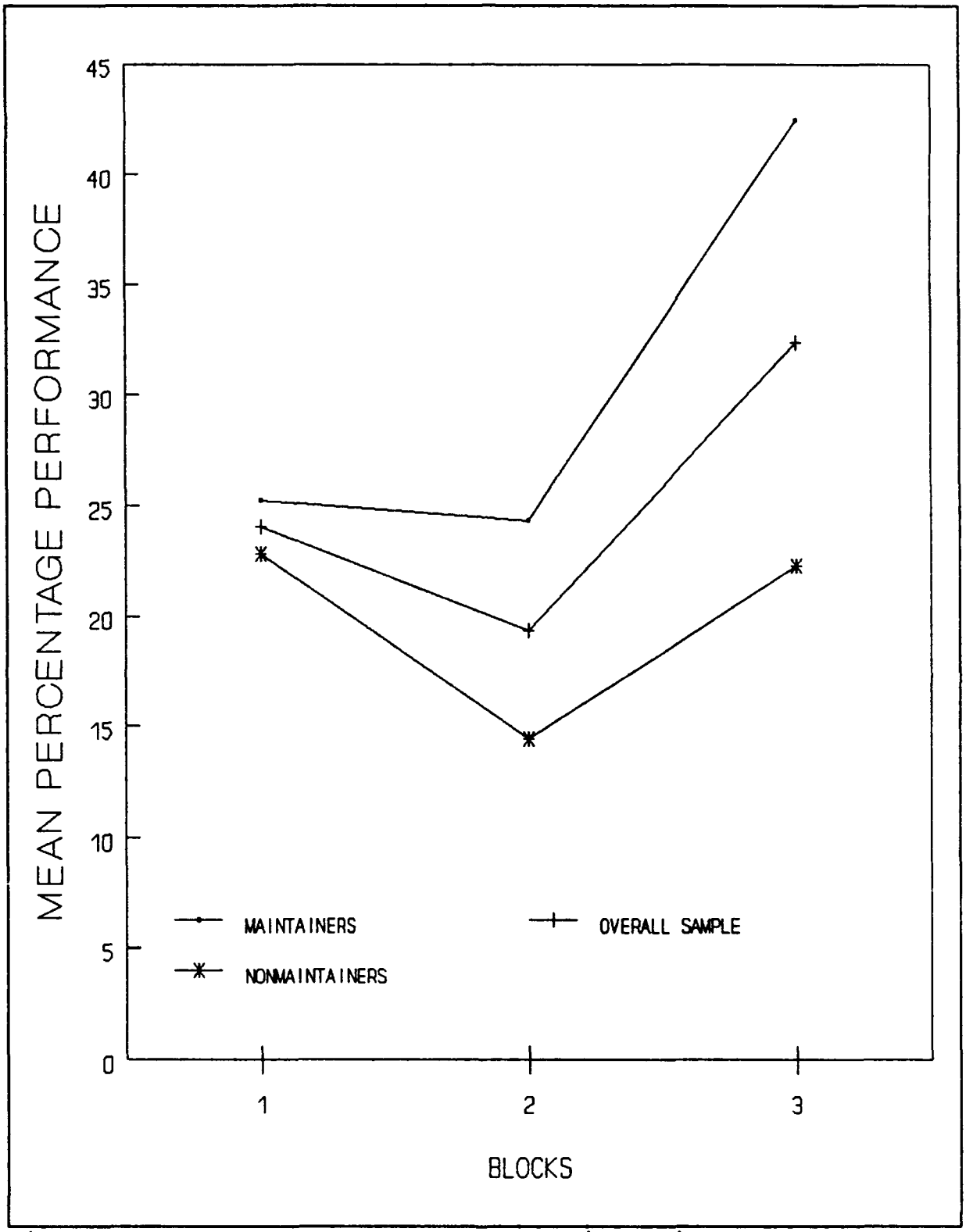

Figure 1. Percentage of correct detections minus false alarms for maintainers, nonmaintainers, and overall sample.

Scale Development

The MMPI was used because it was a readily available pool of items. We did not expect to find a significant relationship between any of the basic MMPI scales and performance since we felt that Army enlistment controls would have severely reduced the number of deviant MMPI scores in our sample (see Table 4). However, in order to determine if the no-relationship assumption was correct, the differences among the means of the validity and clinical subscales of the MMPI for these two groups of subjects were tested by one-way analysis of variance. Five subjects who omitted 30 or more items on the MMPI were not used in computing this analysis in accordance with a recommendation of Greene (1980) who felt that too many omitted items distorted MMPI results. Statistical significance was based on the smallest number of valid cases found in any of the selected variables. None of the means were significantly different.

Mean raw scores with K corrections converted to T-scores are graphically presented in Figure 2. As can be seen, the two groups scored very similarly on each of the subscales, with mean T-scores on both the Schizophrenia and the Hypomania scales moderately elevated. In a related ongoing study, we have found a relationship between the Schizophrenia and Hypomania scales and the likelihood that the soldiers, who participated in the study, successfully completed at least one enlistment (Pierce & Crumley, in preparation).

MMPI items were then submitted to chi-square analyses to determine which items were endorsed differentially by the two groups. Thirty-nine items produced chi-squares with p values less than .05. The MMPI items that discriminated soldiers who were able to maintain adequate performance over the test period from soldiers who showed a performance decrement are presented in Table 5. The direction in which the Performance Maintainers were most likely to respond is indicated at the beginning of each scale item. In Table 6, the direction in which Performance Maintainers were most likely to respond, validity and clinical scale loadings, and the scale deviant response of the selected MMPI items are presented. Even though the MMPI Form R was completed by the soldiers, item number conversions to the Group Booklet Form have been made because "By convention all discussions of individual items in the MMPI literature follows the numbering system of the group booklet form" (Greene, 1980, p.17). Items for which numbers changed are noted in Table 5.

Table 4

Means and Standard Deviations for Validity and Clinical Subscales of the MMPI by Performance Group

	Maintainers		Nonmaintainers	
MMPI Subscale	Mean	Standard Deviation	Mean	Standard Deviation
Cannot Say	2.40	4.85	1.76	3.51
Lie	4.22	1.99	4.50	2.60
F	10.14	7.39	10.18	6.42
K	11.41	4.24	11.24	4.50
Hypochondriasis	13.60	5.22	13.70	4.91
Depression	19.08	5.53	19.63	4.81
Hysteria	19.55	5.49	19.42	6.01
Psychopathic Deviate	25.40	5.19	24.87	4.98
Masculinity-Femininity	24.39	5.24	23.88	5.91
Paranoia	11.61	4.81	11.49	4.46
Psychasthenia	29.35	6.59	28.56	6.44
Schizophrenia	32.67	9.75	32.51	9.68
Hypomania	26.21	4.84	25.87	4.58
Social Introversion	26.60	8.19	26.63	8.72

Note. n=169

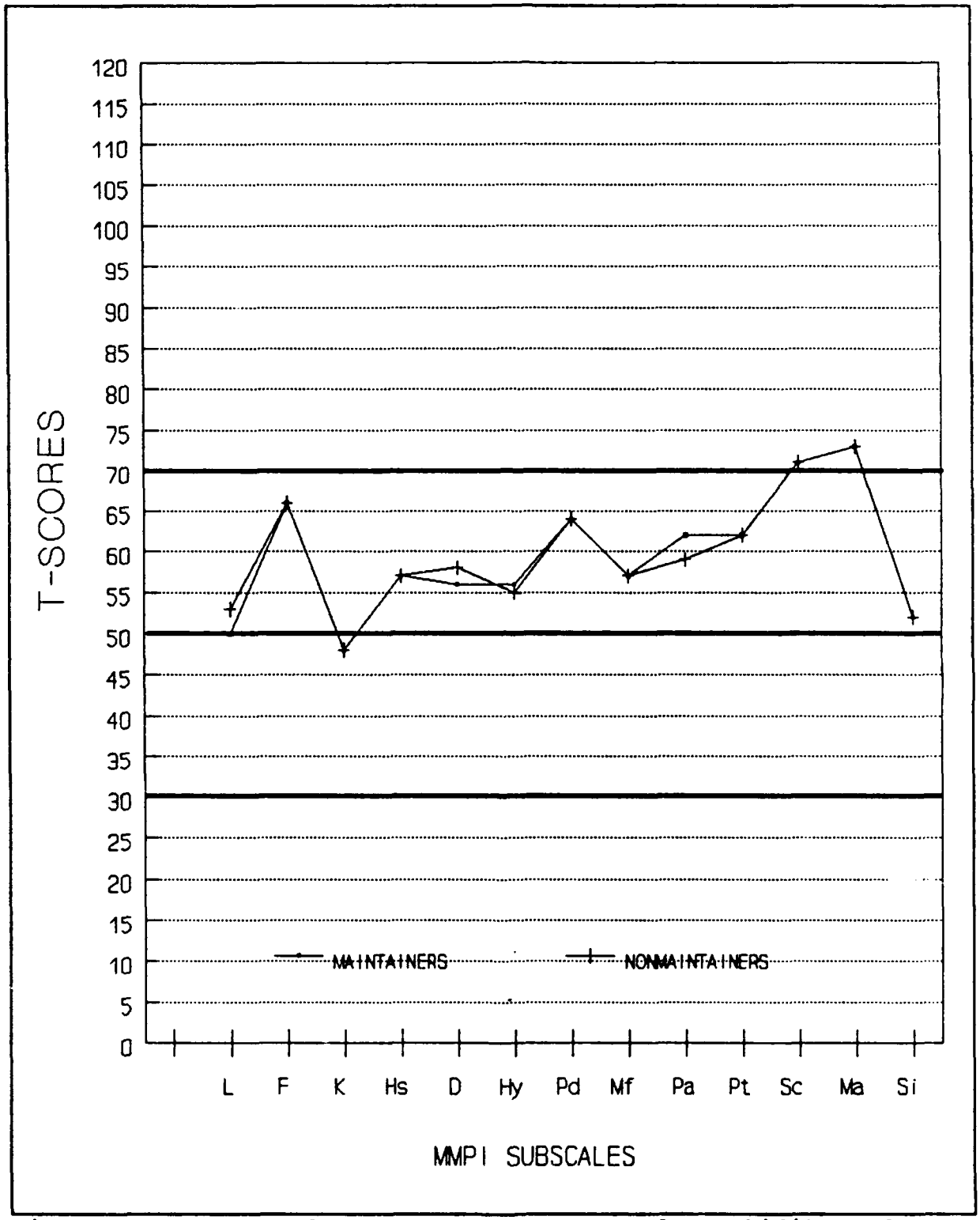

Figure 2. Group performance T-score means for validity and clinical subscales of the MMPI.

Table 5

MMPI Items Found To Discriminate Between Maintainers and Nonmaintainers and the Direction in Which Maintainers Were Most Likely To Respond

True 47. Once a week or oftener I feel suddenly hot all over, without apparent cause.

True 81. I think I would like the kind of work a forest ranger does.

False 84. These days I find it hard not to give up hope of amounting to something.

True 127. I know who is responsible for most of my troubles.

False 136. I commonly wonder what hidden reason another person may have for doing something nice for me.

True 144. I would like to be a soldier.

False 149. I used to keep a diary.

True 150. I would rather win than lose in a game.

False 158. I cry easily.

False 170. What others think of me does not bother me.

True 190. I have very few headaches.

True 210. Everything tastes the same.

False 219. I think I would like the work of a building contractor.

True 245. My parents and family find more fault with me than they should.

False 259. I have difficulty in starting to do things.

True 268. Something exciting will almost always pull me out of it when I am feeling low.

True 270. When I leave home I do not worry about whether the door is locked and the windows closed.

True 274. My eyesight is as good as it has been for years.

False 279. I drink an unusually large amount of water every day.

True 287. I have very few fears compared to my friends.

False 294. I have never been in trouble with the law.

False 298. If several people find themselves in trouble, the best thing for them to do is to agree upon a story and stick to it.

True 302. I have never been in trouble because of my sex behavior.

True 311. During one period when I was a youngster I engaged in petty thievery.

True 320. Many of my dreams are about sex matters.

False 326. At times I have fits of laughing and crying that I cannot control.

*True 371. I am not unusually self-conscious.

*True 376. Policemen are usually honest.

*False 400.If given the chance I could do some things that would be of great benefit to the world.

True 410. I would certainly enjoy beating a crook at his own game.

False 422. I have felt embarrassed over the type of work that one or more members of my family have done.

True 438. There are certain people whom I dislike so much that I am inwardly pleased when they are catching it for something they have done.

False 459. I have one or more bad habits which are so strong that it is no use in fighting against them.

False 467. I often memorize numbers that are not important (such as automobile licenses, etc).

True 497. I enjoy stories of adventure.

True 499. I must admit that I have at times been worried beyond reason over something that really did not matter.

False 510. Dirt frightens or disgusts me.

False 513. I think Lincoln was greater than Washington.

True 535. My mouth feels dry almost all the time.

---

* Item number conversions have been performed and item numbers listed are for the Group Booklet Form. Form R item numbers are shown here in parentheses. 371 (367), 400 (374), 376 (446)

Table 6

The Direction in Which Performance Maintainers (PM) Were Most Likely To Respond, Validity and Clinical Scale Loadings, and the Scale Deviant Response or the Response of Minnesota Normals on the Selected MMPI Items

Item	PM	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si	MN
47	T						T					T			
81	T								F						
84	F							T							
127	T							T		T			T		
136	F						F								
144	T								F						
149	F								T						
150	T	F													
158	F					T				T					
170	F			F			F	F							
190	T				F		F								
210	T		T									T			
219	F								F						
245	T		T					T							
259	F					T						T			
268	T									F			T		
270	T					F									
274	T				F		F								
279	F						F						T		
287	T							F							
294	F							F		F					
298	F												T		
302	T							F				F			
311	T							T				T			
320	T											T			
326	F									T	T	T	T		
371	T													F	
376	t														t
400	F													F	
410	t														f
422	f														f
438	t														f
459	f														f
467	f														f
497	t														t
499	t														f
510	f														f
513	f														t
535	t														f

Note. Table abbreviations are displayed in bold: Validity Scales - Lie, F, K Clinical Scales - Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Masculinity-Femininity, Paranoia, Psychasthenia, Schizophrenia, Hypomania, Social Introversion. For items which are not on any of the primary validity or clinical scales, responses of the majority of Minnesota Normals are shown by lower case letters.

Eleven MMPI items did not appear on any of the three validity scales or on any of the 10 primary clinical scales. Three of these unscored items (376, 497, 513) were answered true by the majority of Minnesota normals, while the remaining eight (410, 422, 438, 459, 467, 499, 510, 535) were answered false by the Minnesota sample. The Performance Maintainers answered in the same way as the Minnesota sample on six of the 11 items. However, the Performance Maintainers answered items 410, 438, 499, 512, and 535 differently than the Minnesota sample.

These 39 items were then used to reclassify the original group of 174 subjects. A discriminant analysis of the responses to this subscale formed a linear function (Wilks's Lambda = .462, Chi Square = 117.96, $p < .0001$) that correctly classified 86% of the cases (88.5% of the Maintainers and 83.9% of the Nonmaintainers).

The final analyses involved determining the amount of improvement in target detection performance which could be expected if the 39-item scale was used for selection. Improvements in target detection performance includes both an increase in targets correctly identified and a decrease in the misidentification of nonexistent targets. As can be seen in Table 7, overall the Nonmaintainers correctly identified an average of 26.17 of the 90 targets presented and, incorrectly reported as targets an average of 135.72 of the 1350 nontarget frames; while the Maintainers correctly identified 34.10 targets and incorrectly reported 104.43 targets. This results in 30.30% more targets being correctly identified by the Maintainers and 29.96% more nonexistent targets being incorrectly identified by the Nonmaintainers.

Since selection of attentive operators will be from the total potential population of operators, it is interesting to determine what improvement can be anticipated when selecting operators whose performance will remain stable over time. In Table 8, the amount of improvement expected if the Maintainers are compared to the entire population is shown. As noted previously, the Maintainers correctly identified 34.10 targets and incorrectly identified 104.43 targets; while the Maintainers and Nonmaintainers, combined, correctly identified 30.13 targets and incorrectly identified 120.09 targets. This results in an increase of 13.18% in targets detected by the Maintainers and 15.00% more nonexistent targets being reported by the entire group.

Table 7

Mean Performance in Target Detection of Maintainers and Nonmaintainers and the Average Percentage Change

	Blocks				Percent Change
	1	2	3	Total	
Correct Detections					
Maintainers	8.45	11.67	13.98	34.10	30.30% (Better)
Nonmaintainers	8.47	8.57	9.13	26.17	
False Alarms					
Maintainers	38.92	31.84	33.69	104.43	
Nonmaintainers	50.30	38.86	46.56	135.72	29.96% (Worse)

Table 8

Mean Performance in Target Detection for the Maintainers and the Total Sample and the Average Percentage Change

	Blocks				Percent Change
	1	2	3	Total	
Correct Detections					
Maintainers	8.45	11.67	13.98	34.10	13.18% (Better)
Total Group	8.46	10.12	11.55	30.13	
False Alarms					
Maintainers	38.92	31.84	33.69	104.43	
Total Group	44.61	35.35	40.13	120.09	15.00% (Worse)

DISCUSSION

In the present study, because of the relatively small sample size of 169, we were unable to retain a portion of the subjects for validation purposes. Thus, this effort should be viewed as a pilot study. The results should be interpreted with caution until validation studies have been completed.

Our primary purpose was to determine if the potential exists for improving target detection performance by creating a scale, based on some parameter of temperament, to identify those who are able to maintain a high level of attentiveness over extended periods of time in situations where there are only occasional rewards. As expected, it was found that individuals do differ in their ability to sustain performance over time, and, as anticipated, we were able to discriminate, based on a subset of MMPI items, between those who did and did not perform in an adequate and consistent manner.

Thirty-nine items were found to discriminate between those who were able to maintain or improve their level of performance across the three 40-minute time periods and those whose performance level was not maintained or improved. This is approximately 7% of the 566 items on the MMPI and because 5% would be expected to discriminate between the two groups merely by chance, these results must be interpreted with caution until validation is completed.

There was, as anticipated, no relationship between any of the MMPI basic validity or clinical scales and performance. However, when the items were studied in an attempt to determine how they related to one another and to define the underlying construct, they appeared to describe a person who was rather conventional; a person with traditionally-defined male interests, and characteristics such as confidence, efficaciousness, and honesty. Future research should help clarify and expand our understanding of the type of person who is likely to perform this type of task most effectively.

We suggest that shortly into a monitoring task, such as the one in the present study, some persons begin to show a fairly significant decrease in the number of targets seen, possibly as a result of less internal discipline. Later into the period--probably as the result of a learning effect followed by an end-spurt--almost all persons show an improvement of performance. However, despite the fact that their performance improves, the group who could not maintain their performance does not improve enough to achieve even their initial performance level. It appears that good attending better is associated with (a) less of an inattentiveness-induced performance decline, and (b) more improvement in searching for and recognizing targets. Obviously, a replication of this study in which subjects are trained on

target detection until they reach asymptote would be of value, since it would provide an estimate of how much of the improvement, in both groups, was the result of learning.

Results of the present study indicate that persons whom we classified as Maintainers detected approximately 30% more targets than did persons classified as Nonmaintainers and when compared to the total population (Maintainers and Nonmaintainers), Maintainers detected approximately 13% more targets. Both groups started with a nearly equal detection rate and both groups improved approximately 50% between the second and third blocks. The difference in targets detected clearly resulted because some subjects performed so badly during the middle one-third of our test. Because of this fall-off, the "end-spurt," which showed improved performance for both groups, could only bring the Nonmaintainers near to their original level, but the Maintainers could improve their initial--and maintained--level by about 50%.

This project is part of a larger program conducted to determine if aptitude, cognitive ability, or temperament could be used to select SSO's who will detect more targets and possibly report fewer false alarms. The result would be an increase in the target acquisition effectiveness of RPV systems. It was determined, in a previous analysis of the data, that the cognitive factor Speed of Closure, particularly as measured by the Snowy Pictures Test, was significantly related to successful target detection (Crumley, Pierce, Schwalm, Coke, & Brown, in preparation). In the present study, no attempt was made to match the groups in terms of ability, as predicted by Speed of Closure tests. However, the difference between Maintainers and Nonmaintainers was not significant during the beginning one-third of the test period (Block 1). We suggest, however, that the relationship between being good at detecting targets because of ability and detecting targets because of good attending behavior is, at best, not highly correlated.

If validation studies support our results, it may be possible to greatly improve performance in target detection by selecting operators who have both (a) the cognitive ability to do well on the task, as evidenced by their performance on the cognitive parameter Speed of Closure, and (b) the temperament necessary to maintain their level of performance over an extended period of time. It appears that some of the potential advantage that accrued to our Performance Maintainers would be lost if only high aptitude operators were used since persons who are already higher level performers have less room to improve. We suggest that this is not the case, since on average only about one-third of the targets were detected. Indeed, some very easy-to-see targets (e.g., a large mess tent in a field) were detected only 80% of the time. Thus, there is much room for improvement.

We conclude that visually detecting targets for an extended period of time from real time terrain presentations is difficult, and that soldiers vary greatly in their ability to perform this type of task. We also suggest that system performance improvements of perhaps one-third could be obtained if SSO's were selected on the basis of cognitive ability, as measured by the cognitive factors tests used by Crumley et al. (in preparation) and, if validation is successful, some version of a temperament scale.

REFERENCES

Barron, F. (1953). An ego-strength scale which predicts response to psychotherapy. Journal of Consulting Psychology, 17, 327-333.

Boff, K.R. and Lincoln, J.E. (1988). Engineering Data Compendium: Human perception and performance. AAMRL, Wright-Patterson AFB, OH.

Brown, D.R., Morgan, W.P., and Kihlstrom, J.F. (1989). Comparison of test construction strategies in an attempt to develop an athletic potential scale. International Journal of Sport Psychology, 20, 93-113.

Butcher, J.N., Jeffrey, T., Cayton, T.G., Colligan, S., DeVore, J.R., and Minegawa, R. (1990). A study of active duty military personnel with the MMPI-2. Military Psychology, 2, 47-61.

Crumley, L.M., Pierce, L.G., Schwalm, R.C., Coke, J.S., and Brown, J. C. (in preparation). Predicting target detection performance using ASVAB subtests and cognitive factor tests. (ARI Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Davies, D.R. and Parasuraman, R. (1982). The psychology of vigilance. London: Academic Press.

Fairbank, J.A., Keane, T.M., and Malloy, P.F. (1983). Some preliminary data on the psychological characteristics of Vietnam veterans with posttraumatic stress disorders. Journal of Consulting and Clinical Psychology, 51, 912-919.

Gayton, W.F., Burchstead, G.N., and Matthews, G.R. (1986). An investigation of the utility of an MMPI posttraumatic stress disorder subscale. Journal of Clinical Psychology, 42, 916-917.

Greene, R.L. (1980). The MMPI: An interpretive manual. Dallas: The Psychological Corporation.

Head, H. (1923). The conception of nervous and mental energy-vigilance: a physiological state of the nervous system. British Journal of Psychology, 14, 126-147.

Keane, T.M., Malloy, P.F., and Fairbank, J.A. (1984). Empirical development of an MMPI Subscale for the Assessment of Combat-Related Posttraumatic Stress Disorder. Journal of Consulting and Clinical Psychology, 52, 888-891.

Mackie, R.R. (1987). Vigilance research - Are we ready for countermeasures? Human Factors, 29, 707-724.

Mackworth, N.H. (1948). The breakdown of vigilance during prolonged visual search. Quarterly Journal of Experimental Psychology, 1, 6-21.

Mackworth, N.H. (1950). Researches on the measurement of human performance (Special Report No. 268). London: Medical Research Council, Her Majesty's Stationary Office.

Mackworth, N.H. (1957). Some factors affecting vigilance. Advancements in Science, 53, 389-393.

Parasuraman, R. (1986). Vigilance, monitoring, and search. In K. R. Boff, L. Kaufman, & J.P. Thomas (Eds.), Handbook of perception and human performance: Vol. II. Cognitive processes and performance. (pp. 43-1-43-39). New York: Wiley.

Pierce, L.G. & Crumley, L.M. (in preparation). Empirical development of a scale to predict completion of first term enlistments. (ARI Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Roland, B., Zelhart, P., and Dubes, R. (1989). MMPI correlates of college women who reported experiencing child/adult sexual contact with father, stepfather, or with other persons. Psychological Reports, 64, 1159-1162.

Sorensen, L.V. and Mors, O. (1988). Presentation of a new MMPI scale to predict outcome after first lumbar diskectomy. Pain, 34, 191-194.

Taylor, J.A. (1951). The relationship of anxiety to the conditioned eyelid response. Journal of Experimental Psychology, 41, 81-92.

Wiener, E.L. (1987). Application of vigilance research: Rare, medium, or well done? Human Factors, 29, 725-736.